An Alternative Approach To Regional Economic Income:
A Fuzzy Logic Model of BEA Economic Areas

A Thesis
Presented to
The Academic Faculty

by

Jamel H. Cato

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in Economics

Georgia Institute of Technology
December 2005

An Alternative Approach To Regional Economic Income:
A Fuzzy Logic Model of BEA Economic Areas

Approved by:

Dr. Willie J. Belton, Jr., Advisor
School of Economics
Georgia Institute of Technology

Dr. Patrick McCarthy
School of Economics
Georgia Institute of Technology

Dr. Usha Nair-Reichert
School of Economics
Georgia Institute of Technology

Date Approved: July 6, 2005

Half the American population no longer reads newspapers: plainly, they are the clever half.

- Gore Vidal

For my mother, Betty Cato Colquitt.
All that I accomplish is because of the foundation you gave me.

ACKNOWLEDGEMENTS

I took the most circuitous route imaginable to get to this point. My graduate coursework and the original version of this thesis were actually completed eleven years ago. The explanation why this thesis is just being submitted now—other than to say that the delay was *not* due to apathy—is a long story better fit for a biography. But here it is and here I am. Incidentally, eleven years ago this thesis would have been filed away in the Library where no one was likely to ever read it again. I am grateful for the advances that now make it available to the world.

I wish to thank my thesis advisor, Dr. Willie J. Belton, Jr., for technical advice, support for my topic and, above all, sticking with me through the most unusual of circumstances. Your experienced guidance helped me craft this thesis into a work of actual relevance for policymakers. I also wish to thank the other knowledgeable members of my thesis committee, Dr. Patrick McCarthy and Dr. Usha Nair-Reichert. I am indirectly indebted to Dr. Bart Kosko, whose fascinating book, *Fuzzy Thinking: The New Science of Fuzzy Logic*, initially sparked my interest in the application of fuzzy logic to economics. Finally, though we've never met, I must also acknowledge Kenneth P. Johnson of the U.S. Department of Commerce for being the intellectual driving force behind BEA Economic Areas. The model proposed in this thesis is a respectful alternative to your impressive approach to the problem of economic unit definition.

While I no longer see t-values in my dreams from the rigor of econometric assignments, the intellectual stimulation I received as a graduate student here will stay with me a lifetime. I acknowledge Georgia Tech for that.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This thesis examines the policy problem of economic unit definition from the perspective of the regional economist. The regional economist faces the challenge of disaggregating macroeconomic activity into subparts that accurately reflect the actual economic organization of a country or region. Such an exercise is important because the Governments of many developed countries rely on it to allocate scarce public resources.

In the United States, the Bureau of Economic Analysis of the U.S. Department of Commerce is responsible for regional economic unit definition. To meet its mandate, the BEA has developed a complex assignment system based principally on commuting flows between regions of the Country. This assignment system works well for the centralized population centers that characterize the majority of the U.S. economy. However, the BEA system is less effective at reflecting the economic organization of rural areas, where there is little interregional commuting. To address this problem, the BEA has developed a practice of using newspaper circulation data as a proxy for economic organization.

In this thesis I develop a partial set model of regional economic organization based on the mathematics of fuzzy logic and propose it as a superior alternative to the BEA's method.

CHAPTER 1

INTRODUCTION

The federal government, through its various units, geographically segregates the Unites States into a variety of metrics – census tracts, zip codes, judicial districts, metropolitan statistical areas, etc. The Bureau of Economic Analysis of the U.S. Commerce Department ("BEA") maintains a special type of metric known as *BEA Economic Areas* ("EA"). BEA Economic Areas were created for the purpose of performing and providing reliable data for regional economic analysis.

The BEA assigns different areas of the country to particular EAs based on a complex three-stage process. A part of this process involves grouping economically related counties into subareas called *Component Economic Areas* ("CEA"), which are then aggregated to form EAs. CEAs are largely based on commuting flows between and within counties, so they inevitably become less reliable at measuring the economic organization of the more remote areas of the country, where very little commuting actually takes place. The BEA has responded to this challenge with a questionable solution: In those regions where no significant commuting data exists, newspaper circulation data is substituted in its place. Referring to this technique in its 1995 redefinition of the EA standard, the BEA wrote:

> "Most of the other counties were preliminarily assigned to nodes on the basis of the locations of the regional newspapers that are most widely read in those counties, according to newspaper circulation data."[1]

---

[1] Johnson, K., "Redefinition of the BEA Economic Areas", Survey of Current Business, Vol. 75, 1995, p. 76

The focus of this thesis is the BEA's process for assigning remote areas to CEAs. In particular, I propose an alternative approach for making such assignments based on the quantitative methods of fuzzy set theory. The central goal of this thesis is to demonstrate that fuzzy logic is the superior method.

This thesis is composed of three principal divisions. Chapter 2 provides necessary background information on fuzzy logic and fuzzy set theory. Chapter 3 focuses exclusively on BEA Economic Areas, including their historical development and the specifics of the CEA assignment process. Finally, in Chapter 4, I present my fuzzy logic model, including a complete elaboration of its principal techniques.

The improper construction of Component Economic Areas, while admittedly arcane, is an important policy problem in at least two regards:

First, various units of the federal government rely on this metric, either directly or indirectly, to make policy and funding decisions. The most prominent example is the Department of Transportation, which directly measures the interregional movement of goods by activity levels between BEA Economic Areas[2]. These measurements are in turned used as a basis for determining the allocation of federal transportation funding.

Second, as the only complete economic time series of regional economic activity, State governments rely on this metric to measure local economic growth and to make policy decisions regarding the allocation of funding for economic development.

---

[2] See "How National Transportation Analysis Regions (NTARs) were defined for the CFS," Bureau of Transportation Statistics, U.S. Department of Transportation, February 29, 1996.

Thus, the inaccurate construction of BEA Economic Areas will directly result in sub-optimal economic policy decisions on the both the national and regional levels.

CHAPTER 2

THEORETICAL REVIEW

2.1 Background Principles

This section contains a basic introduction to Fuzzy Logic and Fuzzy Set Theory. It has been included to provide the background knowledge necessary to fully understand the model set forth in Chapter 4 of this thesis.

Fuzzy Logic is best defined as a form of mathematical logic in which truth can assume a continuum of values between 0 and 1[3]. The notion that every proposition must be either true or false is known as *bivalent logic*[4]. The central idea of fuzzy logic is that every proposition, in addition to being true or false, can also be partially true or partially false. Furthermore, fuzzy logic allows a given proposition to be partially true and partially false at the same time. Fuzzy Set Theory is a system of expressing partial truth mathematically.

Like many theories, Fuzzy Logic is best understood through examples. Take the statement, "Georgia Tech is a large school." Bivalent logic allows this statement to be either true or false. Fuzzy Logic, on the other hand, contends that this statement is 100% true if student enrollment is 10,000 or more, but only 50% true if enrollment is 3,000 and 0% true is enrollment is 500.

---

[2] Definition provided by the WordNet® Database of the Cognitive Science Laboratory of Princeton University, 2003, http://www.cogsci.princeton.edu/~wn/

[3] In Logic, the principle of bivalence is that for any proposition P, either P is true or P is false.

The adjective “fuzzy” is a reference to the fact that Fuzzy Logic attempts to mimic the reasoning process of the human brain in dealing with uncertainty.

2.2 Historical Foundations

Lotfi Zadeh, a professor at the University of California at Berkeley, first introduced the theory of Fuzzy Logic in a 1965 paper entitled, “Fuzzy Sets”[5].

While Zadeh’s paper is rightfully considered the seminal work in the field, it was by no means the first consideration of the nature of truth.

The philosophical foundation of Fuzzy Logic can be divided into four major contributions. The first contribution belongs to Aristotle, who in 200 B.C. proposed the “Law of the Excluded Middle”, which held that every proposition must be either true or false[6]. Plato was among the first to suggest that this dichotomy did not fully describe reality. He theorized that there was a state between true and false. Plato’s observation is generally regarded as the precursor to Fuzzy Logic. In 1920, building on the notions of Plato and others, the logician Jan Lukasiewicz proposed the mathematics for a tri-valued logic which included the concept of fractional truth[7]. He referred to this third logic value (beyond true and false) as “possible”. Lukasiewicz’s ideas formed the basis for a wide body of research into what eventually became known as *many-valued logic*. Modern

---

[5] L. Zadeh, “Fuzzy Sets”, Information and Control, Vol. 8, 1965, pp. 338-353
[6] S. Korner, "Laws of Thought," Encyclopedia of Philosophy, Vol. 4, MacMillan, NY: 1967, pp. 414-417.
[7] C. Lejewski, "Jan Lukasiewicz," Encyclopedia of Philosophy, Vol. 5, MacMillan, NY: 1967, pp. 104-107.

Fuzzy Logic is, in its purest sense, an application of many-valued logic to classic set theory.

2.3 Fuzzy Set Theory

Fuzzy Logic expresses partial truth through the mathematics of set theory. Classic set theory holds that an object is either a member or a nonmember of a given set (but never both). Full membership is indicated by the value 1, while non-membership is indicated by the value 0. Sets of this type are known as *classic sets* or *crisp sets*. Fuzzy logic allows an object to have any value on the continuum between 0 and 1 (including 0 and 1), depending on the degree of membership of said object in a given set. This idea of partial set membership is the cornerstone of Fuzzy Set Theory.

Fuzzy Set Theory is based on Fuzzy Sets. A fuzzy set is a class with no sharp boundary between membership and non-membership. Mathematically, a fuzzy set is a set whose grade of membership falls within the real incursive interval [0,1].

Fuzzy sets exist within a Universe of Discourse, which is the range of all possible values for objects in a given set. If the set is composed of natural language statements, then the Universe of Discourse is the concept that the object's variables describe.

Given a Universe of Discourse X with elements x, a crisp set A of x is defined by the *characteristic function*:

$f_A(x)$ of A, $f_A(x) : X \rightarrow 0,1$	Where, $f_A(x) = 1$ if $x \in A$; $f_A(x) = 0$ if $x \notin A$.

However, the fuzzy set A′ of X (over the same Universe of Discourse) is defined by the *membership function*:

$$\mu_{A'}(x),\ \mu_{A'}(x) : X \rightarrow [0,1]$$

Where,

$\mu_{A'}(x) = 1$ if $x \in A$;

$0 < \mu_{A'}(x) < 1$ if x partially belongs to A′

$\mu_{A'}(x) = 0$ if $x \notin A$.

As you can see, membership functions are generalizations of characteristic functions. They serve an important role in fuzzy set theoretic operations and are discussed in more detail below. For now I would like to further elucidate the concept of fuzzy sets with another example.

Consider the Universe of Discourse consisting of all American economics graduate students, denoted by *X*, and the subset of elite economics graduate students, denoted by *A* (of X). For demonstrative purposes, assume that the set of elite students is composed of students who scored in the 90th percentile or higher on the GRE subject test in economics. The crisp set of elite students would be defined by the characteristic functions: $f_A(x) = 1$ if $x > 90^{th}$ percentile; or $f_A(x) = 0$ if $x < 90^{th}$ percentile. The crisp set would have the following graphic representation:

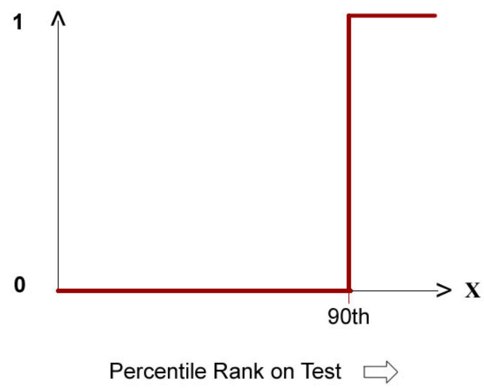

Figure 1
Classic Set of Elite Students

The corresponding fuzzy set of elite students would be defined by the membership functions: $\mu_{A'}(x) = 1$ if $x > 95^{th}$ percentile, $\mu_{A'}(x) = 0$ if $x < 85^{th}$ percentile, $\mu_{A'}(x) = [(x - 100/100]$ if 85^{th} percentile $< x < 95^{th}$ percentile. The fuzzy set would result in a sloped curve of the type shown below.

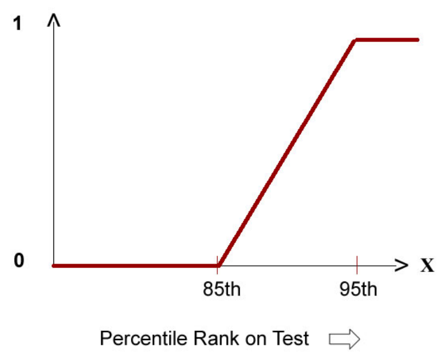

Figure 2
Fuzzy Set of Elite Students

2.4 Fuzzy Set Operations

Here is an overview of selected fuzzy set operations. These concepts are integral to an understanding of the model set forth in Chapter 4.

Table 1
Standard Membership, Containment & Equivalence Functions

Membership, Containment & Equivalence

Membership

Consider **X**, a collection of objects a, b, c

$$\mathbf{X} = \{a, b, c\}$$

A (crisp) set **A** is a collection of some of the objects in **X**, defined as:

For all $x \in \mathbf{X}$ (that is , x = a, b, or c)

either $x \in \mathbf{A}$: x belongs to **A**

or $x \notin \mathbf{A}$: x does not belong to **A**

Containment

$\mathbf{A} \subset \mathbf{B} \Rightarrow$ if $x \in \mathbf{A}$, then $x \vartriangle \mathbf{B}$

Equivalence

$\mathbf{A} = \mathbf{B} \not\supseteq \mathbf{A} \triangleright \mathbf{B}$ and $\mathbf{B} \triangleright \mathbf{A}$

Table 2
Special Sets

Special Sets	
	Null Set ⊆ (empty set) **Universal Set X** (whole set) **Power Set** 2^{X} (set of all subsets of **X**) **Complement** of **A**: if x ▽ **A**, then x △ **A**

Table 3
Basic Set Operations

Basic Set Operations	Operation	Operator	Function (Given Sets A and B)
	Intersection	∩	$\mu_{A}\cap B(x) = \min\{\mu_{A}(x), \mu_{B}(x)\}$, $x \in \mathcal{X}$
	Union	∪	$\mu_{A}\cup B(x) = \min\{\mu_{A}(x), \mu_{B}(x)\}$, $x \in \mathcal{X}$
	Complementation	Ā (for a Set ***A***)	$\mu_{\bar{A}}(x) = 1 - \mu_{A}(x)$, $x \in \mathcal{X}$

2.5 Building a Fuzzy Logic Model

A fuzzy logic model consists of five elements: *Input Variables*, *Linguistic Variables*, *Fuzzy Sets*, a *Rule Base* and a *Final Output*.

Input variables represent the data that the fuzzy system will analyze. They are comprised of crisp numbers or numerical measurements.

A linguistic variable is a variable whose values are words instead of numbers. Linguistic variables describe phenomena that cannot be easily defined in quantitative terms. The unit of measurement for a linguistic variable is a linguistic term. A linguistic term is a descriptive word that modifies a linguistic variable, such as "very" or "somewhat". In a typical fuzzy model, each linguistic variable is attached to a set of linguistic terms that set forth the range of possible values for that variable. Mathematical operations can be performed on these linguistic sets using fuzzy set theory.

A rule base is a group of "if-then" statements that establish relationships among the linguistic variables. Each rule has two parts. The "if" part of the rule is known as the *premise* or *antecedent*, while the "then" Chapter 2s the known as a *consequent*. The actual rules (and the membership degrees associated with them) are typically composed by a human expert based on his or her experience and/or expertise. That is why the rule base is said to contain the intelligence of a fuzzy system.

The Final Output is simply a crisp number or value[8] that represents the result produced by the system.

Now that I have defined the principal elements of a fuzzy model, another example will help clarify their roles. Consider an air conditioner whose fan speed is controlled by a fuzzy system. This air conditioner has a thermostat that constantly measures the temperature in the room, which will always be an exact value.

The rule base of the system consists of the following rules:

- If the temperature is hot, then fan speed is high.
- If the temperature is warm, then fan speed is medium.
- If the temperature is cool, then fan speed is low.

The crisp measurement of the temperature is the input variable. The concept of "temperature" is the model's linguistic variable, with the fuzzy set (hot, warm, cool) serving as its linguistic terms. The final output will be the exact number of revolutions per second the fan will turn based on the fuzzy speed selected by the system.

There are three steps to building a fuzzy logic model: Composition, Inference and Decomposition.

Composition is the process of converting input numbers composed of crisp numbers into fuzzy sets composed of linguistic variables. This is accomplished through the use of membership functions, which map each point in a fuzzy set to the real interval [0,1]. This process is also known as *fuzzification*.

---

[8] While it is theoretically possible to produce a final output that is a fuzzy number or a fuzzy set, fuzzy logic models are rarely designed this way because all types of models seek an optimal result that is best represented by a crisp value.

The next step is to evaluate the fuzzy sets in the model against the rules in the rule base to determine which rules are applicable. In nearly all cases, more than one rule in the rule base will apply. The membership degrees of the respective rules govern which rule is ultimately selected or which course of action is chosen. This process is called *Inference*.

The final step, decomposition, involves converting the fuzzy set chosen in the inference stage of the model into a crisp output value to which we can assign a real world value or meaning. There are dozens of ways to decompose a fuzzy set[9], but the model proposed in Chapter 4 of this thesis employs the *MAX-MIN* method, which evaluates the degree of support for each rule and selects the highest magnitude.

[9] For an accessible discussion of the major defuzzification methods currently in practice, see http://en.wikipedia.org/wiki/Defuzzification

CHAPTER 3

CURRENT PRACTICE

3.1 BEA Economic Areas

The BEA officially defines Economic Areas as follows:

> BEA Economic Areas are a set of Geographic Areas, defined in terms of counties that exhaust the area of the Nation[10].

However, the private sector has developed a number of alternative definitions. Perhaps the most useful and informative is the one provided by the Center for Business & Economic Research at the University of Louisiana at Monroe:

> BEA economic areas are nodal, functional areas defined to facilitate regional economic analysis. Each economic area consists of an economic node--a metropolitan statistical area (MSA) or a similar area that serves as a center of economic activity--and the surrounding counties that are economically related to the center. Commuting patterns are a major factor used in determining the economic relationships among counties, and, to the extent possible, each economic area includes the place of work and place of residence of its labor force. The economic areas cover the entire Nation[11].

[10] As presented on the BEA's official website, which can be found at http://www.bea.gov

[11] Center For Business & Economic Research, University of Louisiana at Monroe, "BEA Regional Projections", Monroe, Louisiana, 1998, pp.1

As a metric, BEA Economic Areas belong to the group of *optional statistical purpose areas* maintained by the federal government. There are two types of SPAs: optional and mandatory. Mandatory SPAs, such as census tracts, are mandated by law, whereas optional SPAs, generally speaking, exist to address specific data needs. EAs were developed to facilitate economic analysis of areas smaller than states but larger than counties.

At present, the BEA recognizes 172 Economic Areas. A complete listing of all Economic Areas and their component counties can be found online at http://www.bea.doc.gov/bea/regional/docs/econlist.asp.

EAs have three primary distinguishing characteristics:

1. They are the only consistently produced economic time series for all local areas;
2. They cover the entire country;
3. They are among the few federal statistical measures that cross state lines.

3.2 Historical Background

The BEA created Economic Areas in 1969. They were developed to meet the data needs of researchers interested in regional economics. They have undergone three major redefinitions since their creation[12]. The first change occurred in 1974, when several new EAs were added to the original set of 167 based on commuting data from the 1960 Census. A similar revision occurred in 1977 based on commuting data from the 1970 Census. The final and most substantial change took place in 1995 when the BEA added

[12] See Johnson, K., "Redefinition of BEA Economic Areas", Survey of Current Business, Vol. 75, 1995

two new steps to its procedure for determining EAs. The first step was the creation of a new metric called a *Component Economic Area*. A CEA is a subarea consisting of one economic node and its surrounding counties. Economic nodes are nearly exclusively composed of Metropolitan Statistical Areas, a mandatory SPA maintained by the Office of Management & Budget. The second new step in the 1995 redefinition was the grouping of CEAs into EAs. Prior to 1995, EAs were defined by the direct grouping of MSAs. There have been no major redefinitions of the EA standard since 1995. In 1996, the Bureau of Transportation Statistics began to define its National Transportation Analysis Regions on the basis of EAs.

3.3 Assignment Process

The method the BEA employs to assign non-metropolitan counties to CEAs is the focus of this thesis.

I chose non-metropolitan areas because the BEA's process for assigning metropolitan counties to CEAs is largely unassailable. Of the nation's 3,141 total counties, 836 are part of existing urban population centers. These 836 counties are easily assigned to CEAs on the basis of their clear economic linkages to the urban centers that compose the economic nodes of those CEAs.

However, assignment of non-metropolitan areas to CEAs is not so straightforward. To start, the majority of non-metropolitan counties are rural and geographically isolated, so there is little commuting data to analyze. Also, the population densities are often too small to be statistically significant.

The BEA has responded to this analytical challenge with a four-part assignment process for non-metropolitan counties. First, the BEA conducts its standard analysis of the limited commuting data available for these counties. Based on this analysis, the BEA has determined that 923 of the nation's 2,035 non-metropolitan counties make up the "hinterlands" of existing urban centers. Second, Of the 1,112 remaining non-metropolitan counties, 130 are the addresses of the most widely circulated newspapers. Third, 68 of this second set have populations of more than 50,000. Finally, 35 of these 68 remaining counties qualify as nodes because they have economic connections to at least five counties.

The weak link in the BEA's system is readily apparent: More than 130 non-metropolitan counties are assigned to CEAs based on the locations of newspapers.

Chapter 4 below presents an alternative approach to making such assignments.

3.4 Alternative Approaches to the BEA Model

BEA Economic Areas (and the alternative composition of them proposed in this thesis) are of course not the only approach to classifying regional economic activity. The BEA's method is to measure economic activity by interregional commuting patterns. There are two primary competing approaches: (1) Classification by industry and (2) Classification by geography.

The industry-based approach is exemplified in the U.S. Census Bureau's Economic Census, which is a "a detailed portrait of the Nation's economy once every five years, from the national to the local level."[13] The Economic Census groups regional economic activity by establishment and industry, based on its North American Industrial Classification System (the successor to the well-known Standard Industrial Classification System). The fundamental difference between the BEA and Census approaches is that the BEA method is a *disaggregation* of national income accounts whereas the Census method is an *aggregation* of firm-level data.

The geography-based approach is embodied in the Office of Management and Budget's statistical area definitions, the most notable of which is its *Metropolitan Statistical Area* metric. Like the BEA, OMB divides the nation into economic regions consisting of central nodes and surrounding areas. The principal difference between the two is that OMB's system is based on geographic distance from population centers, whereas BEA emphasizes economic integration and de-emphasizes geography. Although the OMB approach is clearly more artificial and less dynamic, it is interesting to note that its metric is more widely used.

[13] See "Guide to the 2002 Economic Census", U.S. Census Bureau, U.S. Department of Commerce, 2002 Economic Census.

CHAPTER 4

METHODOLOGY

As I stated in the Introduction, the purpose of this thesis is to propose an alternative method of assigning non-metropolitan counties to Component Economic Areas. The method I propose is a fuzzy logic model.

In designing a fuzzy logic model, like any other model, one must make methodological choices. In my model I have chosen to follow and adapt the approach taken by Lindstrom[14], who designed a fuzzy system to model the impact of interest rates on the level of investment activity in Sweden and also that of Giles[15], who developed a fuzzy approach to measure the underground economy of New Zealand. I selected these particular approaches because they are among the relatively few applications of fuzzy set theory to economic modeling, but also because they contain several important parallels to this thesis. For instance, both works attempt to model an economic region, as I have done here.

4.1 Input Variables

The first step in building my fuzzy system is to select and define the model's base variables. My model contains two variables: Populations Size, which I will denote hereafter as simply "P", and the Media Penetration Rate, which I will denote as "MPR".

---

[14] Lindstrom, T., "A Fuzzy Design of the Willingness to Invest in Sweden", Journal of Economic Behaviour and Organization, Vol. 36, 1998, pp. 1-17
[15] Giles, D.E.A., "Modelling the Hidden Economy and the Tax-Gap in New Zealand", Empirical Economics, Vol. 24, Issue 4, 1999, pp. 621-640.

I considered many different variables and variable combinations before finally settling on the two listed above. In the end, my choices were limited by two constraints. First, in a fuzzy system, as we shall see shortly, the size of the rule base is a direct function of the number of input variables. Generally speaking, each input variable requires at least three input terms (due to the fact that even the simplest fuzzy inference method, the triangular membership function, requires a left boundary, a right boundary and a peak.) In turn, each input term, which also is known as a *linguistic variable*, requires its own *if-then* rule. The relatively straightforward two-variable model set forth in this thesis requires 25 separate rules. So it is easy to see how a more complex system could result in prohibitive rule explosion. Second, I faced the same challenge that BEA analysts do: many potentially powerful variables have no reliable data source.

4.2 Description of Variables

<u>Population Size</u>

My first input variable is simply a measure of each county's residential population based on census counts.

<u>Media Penetration Rate</u>

My second variable is an index that measures the number of television stations whose broadcast signals reach each non-metropolitan county. It is based on the data contained in the Federal Communications Commission's TVQ TV Database[16]. This database is highly useful because it has the ability to produce a specialized map of the broadcast range for any television station in the United States. In this thesis I used that capability to discover how many signals reach each unassigned, non-nodal county. The more signals that reach an area, the higher its MPR index.

I contend that the use of this type of index is superior to the BEA's current approach on at least two grounds.

First, television viewership is a better barometer than newspaper readership for measuring economic association. According to data maintained by the Media Info Center at Northwestern University[17], 90% of all American adults watch television on a daily basis, whereas only 55% read a daily newspaper. The almost universal penetration of television means that, other than the amount of TV viewed, there are very few variations based on demographics. Furthermore, as time passes, newspaper readership is decreasing while television viewership is increasing[18].

Second, the BEA's approach does not account for changes in legislation or media ownership. For example, the BEA completed its last major redefinition of Economic Areas in 1995. A year later, Congress passed the Telecommunications Act of 1996[19], which changed media ownership rules and resulted in changes at several of the newspapers that the BEA relied on to assign non-metropolitan counties to component

---

[16] TVQ TV Database, Federal Communications Commission, Media Bureau, Video Division, http://www.fcc.gov/fcc-bin/audio/tvq.html

[17] Media Management Center, Northwestern University, "Audience Penetration 1940-2003", Evanston, Illinois, 2004.

[18] See Pew Research Center For People and The Press, "The Age of Indifference: A Study of Young Americans and How They View The News", Washington, DC, 1990, pp. 18-20

[19] Telecommunications Act of 1996, Public Law 104-104, February 8, 1996.

economic areas. One of the major weaknesses of the current BEA approach is that it could result in a situation where a CEA is redefined because a newspaper changes its address or merges with another organization, even though the actual economic organization of that region will not have changed. The MPR index, to the contrary, is not as susceptible to this weakness because the FCC database is updated weekly and, for topographical reasons, a change in the ownership of a television station will not typically result in a relocation of a broadcast tower, so, in theory, the MPR index will not be affected.

4.3 Potential Weaknesses

The alternative approach outlined here, while allegedly superior to the BEA's current method, is not without its own potential weaknesses.

There first is the efficacy of Television broadcast signals as a proxy for the economic organization of a community. As such a proxy, broadcast signals suffer from uneven access in the rural communities the BEA covers. Broadcast signals are only an effective metric where they exist. Unfortunately, there are many rural communities in America that receive no television broadcast signals at all. This lack of access is sometimes attributable to topography (broadcast signals are blocked by high mountains), but more often it is the case that broadcasters refuse to provide access because it cannot be economically justified.[20]

[20] See "Advanced Telecommunications in Rural America: The Challenge of Bringing Broadband Service to All Americans", United States Department of Commerce, National

The second potential weakness is the viability of my *Media Penetration Rate* as a substitute for newspaper circulation data. Both mediums are indirect measures of economic activity. Therefore it is possible that neither is an accurate indicator of the economic organization of rural areas.

Although neither of the above limitations is directly applicable to the target area that is the subject of this research, the conclusions I draw should be considered in light of them.

4.4 Target Area

To fully elaborate my model, I will construct a complete fuzzy model for one sample Economic Area. My sample area is BEA Economic Area No. 121, North Platte, Nebraska. I chose this EA because it has the smallest population size of all the Economic Areas and is among the most rural.

The North Platte EA is composed of the following counties and county equivalents:

Telecommunications and Information Administration, United States Department of Agriculture, Rural Development, Rural Utilities Service, April 2000.

Table 4
Counties Within North Platte Economic Area

- Sedgwick, CO
- Arthur, NE
- Blaine, NE
- Chase, NE
- Deuel, NE
- Garden, NE
- Hooker, NE
- Keith, NE
- Lincoln, NE
- Logan, NE
- McPherson, NE
- Perkins, NE
- Thomas, NE

My primary sample period is 1995 – 2000. I chose this period because it takes place after the BEA's major redefinition of Economic Areas in 1995 and incorporates the most recent Census of the Population.

The total population size of the North Platte EA for each year in the sample period was as follows[21]:

Table 5
Population Size By Year for North Platte EA

1995	1996	1997	1998	1999	2000
61,276	61,348	61,345	61,251	61,832	61,707

[21] According to data maintained by the U.S. Census Bureau at http://www.census.gov

The corresponding Media Penetration Rate for the North Platte EA for each year in the sample period was as follows[22]:

Table 6
MPR by Year for North Platte EA

1995	1996	1997	1998	1999	2000
0.00508	0.00507	0.00507	0.00509	0.00505	0.00505

4.5 Composition

To create fuzzy sets, we first need to define our input variables, which you will recall are ***P*** and ***MPR***. For each input variable, I defined five linguistic variables (Low, Somewhat Low, Normal, Somewhat High, and High). The goal of the linguistic variables is to derive - for each year in the series – an *output variable*, which in this case is a measure of this EA's suitability as a node of a component economic area. I will refer to our output variable as a Nodal Index, and denote it as ***NI***. The Nodal Index has its own set of linguistic variables, which I have defined as Low, Medium, and High.

To help clarify what each linguistic variables means, I established benchmarks by setting the value of the "normal" linguistic variable as the five-year simple moving average of the sample data. So, for example, if I were to say that *P* or *MPR* for the year 1997 was "normal", it would be a reference to the average of the period 1992 – 1997, and for 1998 it would be a reference to the average of the period 1993 – 1998 and so on.

[22] The calculate *MPR*, I divided the total population for each EA by the number of stations whose broadcast signals reach the EA. The figures listed in Table 3 are the quotients of those calculations for each respective year.

Having established "normal" values, I can assign a degree of membership for each input variable for each year. I achieved this by adopting Lindstrom's method of setting each linguistic variable as either one or two standard deviations from the normal value, based on the distance of the variable from the normal value. The formula I used for this is:

$$s_N = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^2}$$

which means that for any given year the degree of support for either input variable would be:

Table 7
Standard Deviations From Mean For Linguistic Variables

"Low"	**"Somewhat Low"**	**"Normal"**	**"Somewhat High"**	**"High"**
(L)	(SL)	(N)	(SH)	(H)
-2 SD_t	-1 SD_t	$Mean_t$	+1 SD_t	+2 SD_t

Where,

SD = standard deviation

t = sample time period

The result is two five-number fuzzy sets for each input variable for each year. These sets are known as *breakpoints*. For example, the breakpoints for the input variable ***P*** for the year 2000 are:

Table 8
Breakpoints for "Population" Input Variable

(L)	(SL)	(N)	(SH)	(H)
60,968	61,214	61,460	61,706	61,952

The value associated with the linguistic variable (N), 61,460, is the mean for the moving time period (1995-2000), whereas the value associated with the variable (SL), 61,214, is the mean minus two times the standard deviation, which in this case is 246.05.

I then created membership functions for the fuzzy sets. Continuing with the year 2000 example for the variable P, you will note that the actual population was 61,707, which placed it between "somewhat high" and "high" on the scale derived above. Classic set theory dictates that the population is either somewhat high or high. But fuzzy logic, if you recall from the background principles section, allows the figure to be considered both at the same time. That is, the figure is a member of both sets to some degree. The degree of membership in each set corresponds to the distance of the value from each set's breakpoints. The mathematical relationship between the breakpoints and the values are called *membership functions*. Essentially, a membership function assigns a weight to an input value in order to produce an output value.

In this model I have used the most common form, a simple triangular membership function of the type:

$$\mu_{triangle}(x) = \begin{cases} 0 & x < L \\ 1 - \dfrac{|C - x|}{\frac{(R-L)}{2}} & L < x < R \\ 0 & x > R \end{cases}$$

where,

L = the left boundary of the triangle

R = the right boundary of the triangle

C = the peak of the triangle

Here I have assigned weights to the input variables by taking the inverse of the distances between the values. Returning to my Year 2000 example for ***P***, it is evident that the population figure 61,707, while still a member of both sets, is much closer to the "somewhat high" set than it is to "high" set.   In fact, it exceeds the "somewhat high" breakpoint by only one person. Taking the inverse of the standard deviation that I calculated earlier results in the following outcome:

Table 9
Inverse of Standard Deviation For "Population" Input Variable

(L)	(SL)	(N)	(SH)	(H)
0.00	0.00	0.00	0.99593	0.00406

This outcome can be interpreted to mean that a population size of 61,707 is a member of the somewhat high set to a 99% degree and a member of the high set to a 1% degree. This makes sense on a logical basis and a mathematical one because the two membership degrees sum to unity. That is, 0.995935785 + 0.004064215 = 1

This outcome would be described graphically as:

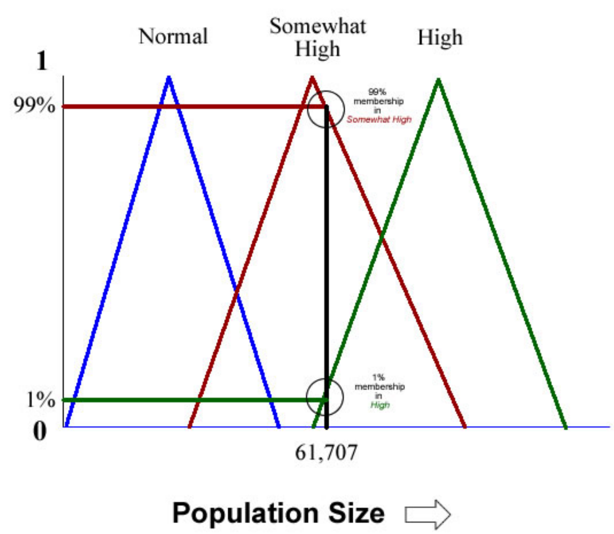

Figure 3
Fuzzy Sets of the Variable *P* for the Year 2000

The next step is to create a rule block. A rule block determines how particular degrees of membership for each input variable combine to produce a single output variable. Based on our model of two input variables with five linguistic variables each and one output variable with three linguistic variables, our rule block will have 25 rules, as reflected in the following table:

Table 10
Rule Block for Cato Model

	IF		**THEN**	
Rule No. ↓	**P**	**MPR**	**NI**	**D.O.S.**
1	H	H	H	1.0
2	H	SH	H	0.8.
3	H	N	M	0.8
4	H	SL	M	0.8
5	H	L	M	0.8
6	SH	H	H	1.0
7	SH	SH	H	1.0
8	SH	N	H	1.0
9	SH	SL	M	1.0
10	SH	L	L	0.8
11	N	H	M	0.8
12	N	SH	M	1.0
13	N	N	L	0.8
14	N	SL	M	0.8
15	N	L	L	0.8
16	SL	H	L	1.0
17	SL	SH	M	0.8
18	SL	N	M	0.8
19	SL	SL	L	1.0
20	SL	L	L	1.0
21	L	H	M	0.8
22	L	SH	M	0.8
23	L	N	M	0.8
24	L	SL	L	1.0
25	L	L	L	1.0

A few explanatory comments about this rule block are in order. First, the appropriate way to interpret the individual rules is to use an "if-then" construct. For example, Rule No. 2 should be understood to say, "If the Population Size is High AND the Media Penetration Rate is Somewhat High, then the Nodal Index is High." Second, the figures in the D.O.S. column represent the degree of validity of each rule, and **not** the degree of membership of a rule in a set. A higher degree of support represents greater rule validity and a lower degree of support represent less validity. So while I have a high degree of confidence in Rule No. 2, I am not as confident in Rule No. 5, which states, "If the Population Size is High AND the Media Penetration Rate is Low, then the Nodal Index is Medium." A large population with a low media penetration rate could reflect the fact that the target area has low income-per-capita, in which case it is suitable to serve as an economic node. On the other hand, a large population with a low media penetration rate could mean the target area is geographically isolated, in which case it would be less suitable to be an economic node (since the BEA bases it definition of economic association on commuting flows). Because I am less certain of the validity of the rule, it has a lower degree of support.

There is no universal method of assigning degrees of support. It is often the case, as it is here, that degrees of support are simply a reflection of modeler's belief system.

Keeping in mind that the purpose of this exercise is to derive a Nodal Index for the North Platte Economic Area, my final task is to convert the linguistic output of the rule block into a "crisp" output number. In other words, the rule block I outlined above will always produce a linguistic output of the form: "If the Population Size is X and the Media Penetration Rate is Y, then the Nodal Index is Z", where Z will be Low, Medium or High. In this step we want to convert “Z” into a number to which we can assign meaning. This conversion process is known as *decomposition*.

4.6 Decomposition

There are dozens of decomposition methods, but here I elected to follow the MIN-MAX approach advocated by Giles. There are two steps in this approach. First, the linguistic variables for NI (Low, Medium, High) are attached to the weighted values 0.25, 0.50 and 0.75. Second, recalling that for each observation of *P* and *MPR*, there are at most two linguistic terms, there are a maximum of four active rules for each NI value.

Returning to my year 2000 sample, the corresponding values for each input variable are:

Table 11
Decomposition Values for Input Variables

P	**(SH)**	**(H)**
	0.99593	**0.00406**
MPR	**(N)**	**(SH)**
	0.5068	**0.5085**

The decomposed values from Table 8 result in the following table of inference values:

:

Table 12
Inference Values for Rule Block Combinations

Combination	Rule No.	NI Intermediate Value	NI Degree
(SH), (H)	6	(H): 1.0 x 0.00406 =	0.00406
(N), (SH)	12	(M): 0.8 x 0.50680 =	0.40544
(H), (N)	3	(M): 0.8 x 0.00406 =	0.00324
(SH), (SH)	7	(H): 1.0 x 0.99593 =	0.99593

Where,

NI Level = MIN (P, MPR)

NI Degree = MAX (P, MPR)

Simply stated, where the inference process produces multiple values, the MIN operator requires me to select the smaller value, while the MAX operator requires me to select the larger value.

Table 13
Weighted Min-Max Values for Cato Model

Linguistic Value	NI Intermediate Value	Weight
(H)	0.00406	0.25
(N)	0.40544	0.50
(SH)	0.99593	0.75

Using the *weighted average decomposition method*, the final Nodal Index can be calculated with the following formula:

$$NI = \frac{\sum_{i=1}^{n} v^i w_i}{\sum_{i=1}^{n} v^i}$$

Where,

NI = Nodal Index

v^i = NI Intermediate Value

w_i = weighted value of rule

Applying this formula to the output of my model produces the following equation:

$$\frac{(0.00406 \text{ x } 0.25) + (0.40544 \text{ x } 0.50) + (0.99593 \text{ x } 0.75)}{(0.00406 + 0.40544 + 0.9593)} = \mathbf{0.694537186}$$

CHAPTER 5

CONCLUSION

The final Nodal Index value of 0.694537186 can be interpreted to mean that, in the year 2000, on the basis of population size and media penetration, the counties which compose the North Platte Economic Area form an appropriate reflection of the region's economic organization. Conversely, a final index of less than 0.50, in my view, would indicate an inaccurate reflection of the region's actual economic organization.

It should be noted that the final Nodal Index produced herein is an alternative approach to the assignment of non-metropolitan counties to component economic areas, and not a wholesale replacement of the BEA model of economic unit definition. With that clarification in mind, I contend that the approach set forth in the thesis is a better method than the current one currently practiced by the Government.

The model set forth in this thesis is valid only for BEA Economic Area No. 121. Beyond that, it should be considered only a conceptual framework; a true test of this method's superiority over the BEA's current method would require a separate fuzzy logic model of all 172 BEA Economic Areas.

REFERENCES

Adams, John S., Van Drasek, Barbara, and Phillips, Eric., Metropolitan Area Definition in the United States, Urban Geography, Vol. 20, 1999, pp. 695-726.

Federal Communications Commission, Media Bureau, Video Division, TVQ TV Database, http://www.fcc.gov/fcc-bin/audio/tvq.html

Johnson, Kenneth P., Redefinition of the BEA Economic Areas, Survey of Current Business, Vol. 75, 1995, pp. 75-81

Kaufmann, A. and M. M. Gupta, *Introduction to Fuzzy Arithmetic Theory and Application*, Van Nostrand Reinhold, New York, 1991

Klir, G.J. and Folger, T.A., *Fuzzy Sets, Uncertainty, and Information*, Prentice Hall, Englewood Cliffs, NJ, 1988

Kosko, B., *Fuzzy Thinking: The New Science of Fuzzy Logic*, Hyperion: New York, 1993.

Kosko, Bart, & Isaka, Satoru, Fuzzy Logic, Scientific American, July 1993, pp. 62--67.

Novak, V., *Fuzzy Sets and Their Applications*, Adam Hilger, Philadelphia, 1987
Pease, B., What's All This Fuzzy Logic Stuff, Anyhow?, Electronic Design, May 1993, pp. 77-79

Rain, David R., Commuting Directionality, a Functional Measure for Metropolitan and Nonmetropolitan Area Standards, Urban Geography, Vol. 20, 1999, pp. 749-767

Richardson, G., The Structure of Fuzzy Preferences: Social Choice Implications, Social Choice and Welfare, Vol. 15, 1998, pp.359-369

U.S. Census Bureau, Population Division, 2000 County-to-County Worker Flow Files, http://www.census.gov/population/www/socdemo/journey.html (Link No. 1)

U.S. Census Bureau, Population Division, Journey to Work and Place of Work Database, http://www.census.gov/population/www/socdemo/journey.html (Link No. 4)

U.S. Department of Commerce, Bureau of Economic Analysis, BEA Economic Areas Component County List, http://www.bea.doc.gov/bea/regional/docs/econlist.asp

U.S. Department of Commerce, Bureau of Economic Analysis, BEA Local Area Personal Income Database, http://www.bea.doc.gov/bea/regional/

U.S. Department of Commerce, Bureau of Economic Analysis, Local Area Web Application, Regional Economic Information System (REIS) CD–ROM 1969–2002, http://www.bea.doc.gov/bea/regional/reis/

U.S. Office of Management and Budget, Alternative Approaches to Defining Metropolitan and Nonmetropolitan Areas Notice, Federal Register, Vol. 63, No. 244, 1998

Wheeler, J., Aoyama, Y., and Warf, B. (eds), *Cities in the Telecommunications Age: The Fracturing of Geographies*, Routledge, New York, 2000

Zimmermann, H.J., *Fuzzy Set Theory and Its Applications*, 3rd ed., Kluwer Academic Publishers, Boston, 1996